ELSEVIER

Available online at www.sciencedirect.com

SciVerse ScienceDirect

Advances in Mathematics 231 (2012) 74–101

ADVANCES IN Mathematics

www.elsevier.com/locate/aim

A central limit theorem for the determinant of a Wigner matrix

Terence Tao[a,*], Van Vu[b]

[a] Department of Mathematics, UCLA, Los Angeles CA 90095-1555, United States
[b] Department of Mathematics, Yale University, New Haven, CT 06520, United States

Received 28 November 2011; accepted 8 May 2012
Available online 1 June 2012

Communicated by the Managing Editors of AIM

Abstract

We establish a central limit theorem for the log-determinant $\log|\det(M_n)|$ of a Wigner matrix M_n, under the assumption of four matching moments with either the GUE or GOE ensemble. More specifically, we show that this log-determinant is asymptotically distributed like $N(\log\sqrt{n!} - \frac{1}{2}\log n, \frac{1}{2}\log n)_{\mathbb{R}}$ when one matches moments with GUE, and $N(\log\sqrt{n!} - \frac{1}{4}\log n, \frac{1}{4}\log n)_{\mathbb{R}}$ when one matches moments with GOE.

Keywords: Random matrices; Determinants; GUE; GOE

Contents

1. Introduction 75
 1.1. Notation 80
2. The central limit theorem for GUE 80
 2.1. The GOE case 88
3. Resolvent swapping: a deterministic analysis 88
4. Proof of Theorem 5 92
 Acknowledgments 96
 Appendix. Moment calculations 97
 References 100

* Corresponding author.
E-mail addresses: tao@math.ucla.edu (T. Tao), van.vu@yale.edu (V. Vu).

0001-8708/$ - see front matter
doi:10.1016/j.aim.2012.05.006

1. Introduction

Random matrix theory is an important subject in mathematics with applications to various areas such as numerical analysis, mathematical physics, statistics, number theory and computer science, to mention a few. One of the main goals of this theory, by and large, is to understand the distribution of various interesting functionals of a random matrix that naturally arise from linear algebra.

One of most natural and important matrix functionals is the *determinant*. As such, the study of determinants of random matrices has a long and rich history. The earlier papers on this study focused on the determinant $\det A_n$ of the non-Hermitian i.i.d. model A_n, where the entries ζ_{ij} of the matrix were independent random variables with mean 0 and variance 1. The earliest paper we find here belongs to Szekeres and Turán [31], in which they studied an extremal problem. In the 1950s, there were a series of papers [12,27,39,28] devoted to the computation of moments of fixed orders of the determinant (see also [15]). The explicit formula for higher moments get very complicated and in general not available, except in cases when the atom variables have some special distribution (see, for instance [8]).

One can use the estimate for the moments and the Chebyshev inequality to obtain an upper bound on the magnitude $|\det A_n|$ of the determinant. However, no lower bound was known for a long time. In particular, Erdős asked whether $\det A_n$ is non-zero with probability tending to one. In 1967, Komlós [21,22] addressed this question, proving that asymptotically almost surely $|\det A_n| > 0$ for random Bernoulli matrices (where the atom variables are i.i.d. Bernoulli, taking values ± 1 with probability $1/2$). His method also works for much more general models. Following [21], the upper bound on the probability that $\det A_n = 0$ has been improved in [18,32, 33,3]. However, these results do not say much about the value of $|\det A_n|$ itself.

A few years ago, the authors [32] managed to prove that for Bernoulli random matrices, with probability tending to one (as n tends to infinity)

$$\sqrt{n!}\exp(-c\sqrt{n\log n}) \le |\det A_n| \le \sqrt{n!}\omega(n) \tag{1}$$

for any function $\omega(n)$ tending to infinity with n. This shows that asymptotically almost surely, $\log|\det A_n|$ is $(\frac{1}{2} + o(1))n\log n$, but does not otherwise provide much information on the limiting distribution of the log determinant. For related works concerning other models of random matrices, we refer to [30].

In [17], Goodman considered random Gaussian matrices $A_n = (\zeta_{ij})_{1\le i,j\le n}$ where the atom variables ζ_{ij} are i.i.d. standard real Gaussian variables, $\zeta_{ij} \equiv N(0,1)_{\mathbb{R}}$. He noticed that in this case the square of the determinant can be expressed as the product of independent chi-square variables. Therefore, its logarithm is the sum of independent variables and thus one expects a central limit theorem to hold. In fact, using properties of the chi-square distribution, it is not hard to prove[1]

$$\frac{\log(|\det A_n|) - \frac{1}{2}\log n! + \frac{1}{2}\log n}{\sqrt{\frac{1}{2}\log n}} \to N(0,1)_{\mathbb{R}}, \tag{2}$$

where $N(0,1)_{\mathbb{R}}$ denotes the law of the real Gaussian with mean 0 and variance 1; see e.g. [29] for a proof. Informally, we may write this law as

[1] Here and in the sequel, $\to$ denotes convergence in distribution.

$$|\det A_n| \approx n^{-1/2}\sqrt{n!}\exp\left(N\left(0, \frac{1}{2}\log n\right)_{\mathbb{R}}\right). \tag{3}$$

We remark that because the second moment of $\exp(N(0,t)_{\mathbb{R}})$ is e^{2t} for any $t > 0$, this law is consistent with the second moment identity

$$\mathbf{E}|\det A_n|^2 = n!, \tag{4}$$

for i.i.d. matrices (and in particular, for Gaussian matrices) that was first observed by Turán [39], and easily derivable from the Leibniz expansion

$$\det A_n = \sum_{\sigma\in S_n} \operatorname{sgn}(\sigma)\prod_{i=1}^{n}\zeta_{i\sigma(i)} \tag{5}$$

after observing that the terms on the right-hand side are pairwise uncorrelated in the i.i.d. case.

A similar analysis (but with the real chi distribution replaced by a complex chi distribution) also works for complex Gaussian matrices, in which ζ_{ij} remain jointly independent but now have the distribution of the complex Gaussian $N(0,1)_{\mathbb{C}}$ (or equivalently, the real and imaginary parts of ζ_{ij} are independent and have the distribution of $N(0,\frac{1}{2})_{\mathbb{R}}$). In that case, one has a slightly different law

$$\frac{\log(|\det A_n|) - \frac{1}{2}\log n! + \frac{1}{4}\log n}{\sqrt{\frac{1}{4}\log n}} \to N(0,1)_{\mathbb{R}}, \tag{6}$$

or more informally

$$|\det A_n| \approx n^{-1/4}\sqrt{n!}\exp\left(N\left(0, \frac{1}{4}\log n\right)_{\mathbb{R}}\right). \tag{7}$$

Again, this remains consistent with (4).

We turn now to more general real i.i.d. matrices, in which the ζ_{ij} are jointly independent and real with mean zero and variance one. In [14], Girko stated that (2) holds for such random matrices under the additional assumption that the fourth moment of the atom variables is 3. Twenty years later, he claimed a much stronger result which replaced the above assumption by the assumption that the atom variables have bounded $(4+\delta)$-th moment [16]. However, there are several points which are not clear in these papers. Recently, Nguyen and the second author [26] gave a new proof for (2) under an exponential decay hypothesis on the entries. Their approach also results in an estimate for the rate of convergence and is easily extended to handle to complex case.

The analysis of the above random determinants relies crucially on the fact that the rows of the matrix are jointly independent. This independence no longer holds for Hermitian random matrix models, which makes the analysis of determinants of Hermitian random matrices more challenging. The Hermitian version of Komlos' result [21,22] was posed as an open question by Weiss in the 1980s and was solved only five years ago [6] and for this purpose the authors needed to introduce the quadratic analogue of Littlewood–Offord–Erdős theorem. The analogue of (1) was first proved in [35, Theorem 31], as a corollary[2] of the Four Moment theorem. But much as

[2] This theorem requires the atom variable has vanishing third moment, but one can remove this requirement using very recent estimates of Nguyen [25] and Vershynin [40] on the least singular value.

in the situation in the non-Hermitian case, these proofs do not reveal much information about the limiting distribution of the determinant.

Let us now narrow down our consideration to the following class of random matrices.

Definition 1 (*Wigner Matrices*)**.** Let $n \geq 1$ be an integer. An *$n \times n$ Wigner Hermitian matrix M_n* is defined to be a random Hermitian $n \times n$ matrix M_n with upper triangular complex entries ζ_{ij} and diagonal real entries ζ_{ii} $(1 \leq i \leq n)$ jointly independent, with mean zero and variance one for $1 \leq i < j \leq n$, and mean zero and variance σ^2 for $1 \leq i \leq n$ and some $\sigma^2 > 0$ independent of n. We refer to the distributions of the ζ_{ij} as the *atom distributions* of M_n.

We say that the Wigner matrix ensemble *obeys Condition* **C1** for some constant C_0 if one has

$$\mathbf{E}|\zeta_{ij}|^{C_0} \leq C_1$$

for all $1 \leq i, j \leq n$ and some constant C_1 independent of n.

Example 2. The famous *Gaussian Unitary Ensemble* (GUE) is the special case of the Wigner ensemble in which the atom distributions ζ_{ij} are given by the complex Gaussian $N(0,1)_{\mathbb{C}}$ for $1 \leq i < j \leq n$ and the real Gaussian $N(0,1)_{\mathbb{R}}$ for $1 \leq i = j \leq n$, thus in this case $\sigma^2 = 1$. At the opposite extreme, the *complex Hermitian Bernoulli ensemble* is an example of a discrete Wigner ensemble in which the atom distributions ζ_{ij} is equal to $\pm\frac{1}{\sqrt{2}} \pm \frac{1}{\sqrt{2}}\sqrt{-1}$ (with independent and uniform Bernoulli signs) for $1 \leq i < j \leq n$, and equal to ± 1 for $1 \leq i = j \leq n$ (so again $\sigma^2 = 1$).

Another important example is the *Gaussian Orthogonal Ensemble* (GOE) in which the atom distributions ζ_{ij} are given by $N(0,1)_{\mathbb{R}}$ for $1 \leq i < j \leq n$ and $N(0,2)_{\mathbb{R}}$ for $1 \leq i = j \leq n$, thus $\sigma^2 = 2$ in this case. Finally, the *symmetric Bernoulli ensemble* is an example in which $\zeta_{ij} \equiv \pm 1$ for all $1 \leq i \leq j \leq n$, with $\sigma^2 = 1$.

All of the above examples obey Condition **C1** for arbitrary C_0.

We now consider the distribution of the determinant for Wigner matrices. We first make the observation that the first and second moments of the determinant are slightly different in the Wigner case than in the i.i.d. case:

Theorem 3 (*First and Second Moment*)**.** *Let M_n be a Wigner matrix.*

- *(First moment). When n is odd, then* $\mathbf{E}\det M_n = 0$. *When n is even, one has*

  $$\mathbf{E}\det M_n = (-1)^{n/2}\frac{n!}{(n/2)!2^{n/2}}.$$

  In particular, by Stirling's formula one has

  $$\mathbf{E}\det M_n = \left(\frac{(-1)^{n/2}2^{1/4}}{\pi^{1/4}} + o(1)\right) n^{-1/4}\sqrt{n!}.$$

- *(Second moment). If M_n is drawn from GOE, then has*[3]

  $$n^{3/2}n! \ll \mathbf{E}|\det M_n|^2 \ll n^{3/2}n!,$$

  while if M_n is instead drawn from GUE, then

  $$n^{1/2}n! \ll \mathbf{E}|\det M_n|^2 \ll n^{1/2}n!$$

[3] See Section 1.1 for the asymptotic notation we will use, including Vinogradov's notation $\ll$.

Proof. See Appendix. More precise asymptotics for these moments in the GUE case were established by Brezin and Hikami [4] (see also [11,13,23]), however we give an elementary and self-contained proof of the above results in the Appendix. □

Even in the GUE case, it is highly non-trivial to prove an analogue of the central limit theorem (6); this was first achieved in [7] via a lengthy computation using the explicit formula for the joint distribution of the eigenvalues. Notice that the observation of Goodman does not apply due to the dependence between the rows and so it is not even clear why a central limit theorem must hold for the log-determinant.

While it does not seem to be possible to express the log-determinant of GUE or GOE as a sum of independent random variables, in this paper we present a way to approximate the log-determinant as a sum of weakly dependent terms, based on[4] analysing a tridiagonal form of both GUE and GOE due to Trotter [38]. Using stochastic calculus and the martingale central limit theorem (see Section 2), we give a new proof of the following result.

Theorem 4 (*Central Limit Theorem for Log-Determinant of GUE and GOE*)**.** *Let M_n be drawn from GUE. Then*

$$\frac{\log|\det(M_n)| - \frac{1}{2}\log n! + \frac{1}{4}\log n}{\sqrt{\frac{1}{2}\log n}} \to N(0,1)_{\mathbb{R}}.$$

Similarly, if M_n is drawn from GOE rather than GUE, one has

$$\frac{\log|\det(M_n)| - \frac{1}{2}\log n! + \frac{1}{4}\log n}{\sqrt{\log n}} \to N(0,1)_{\mathbb{R}}.$$

Informally, this theorem asserts that

$$|\det M_n| \approx n^{-1/4}\sqrt{n!}\exp\left(N\left(0, \frac{1}{2}\log n\right)_{\mathbb{R}}\right)$$

for GUE, and

$$|\det M_n| \approx n^{-1/4}\sqrt{n!}\exp(N(0,\log n)_{\mathbb{R}})$$

for GOE (compare with (3), (7)). Note also that these distributions are consistent with the moment computations in Theorem 3.

As mentioned previously, Theorem 4 has also been proven (using the explicit joint density distribution of the GUE and GOE eigenvalues) by Delannay and Le Caer [7]. However our approach is quite different in nature and somewhat less computational, and may be of independent interest.

The next task is to extend beyond the GUE or GOE case. Our main tool for this is a four moment theorem for log-determinants of Wigner matrices, analogous to the four moment theorems for eigenvalues [35,34,36], Green's functions [10], and eigenvectors [37,20]. Let us say that two Wigner matrices $M_n = (\zeta_{ij})_{1\le i,j\le n}$ and $M'_n = (\zeta'_{ij})_{1\le i,j\le n}$ *match to order m* off the diagonal and to order k on the diagonal if one has

$$\mathbf{E}(\mathrm{Re}\zeta_{ij})^a(\mathrm{Im}\zeta_{ij})^b = \mathbf{E}(\mathrm{Re}\zeta'_{ij})^a(\mathrm{Im}\zeta'_{ij})^b$$

for all $1 \le i \le j \le n$ and natural numbers $a, b \ge 0$ with $a + b \le m$ for $i < j$ and $a + b \le k$ for $i = j$.

[4] We would like to thank R. Killip for suggesting the use of Trotter's form.

Theorem 5 (*Four Moment Theorem for Determinant*)**.** *Let M_n, M_n' be Wigner matrices whose atom distributions have independent real and imaginary parts that match to fourth order off the diagonal and to second order on the diagonal, are bounded in magnitude by n^{c_0} for some sufficiently small but fixed $c_0 > 0$, and are supported on at least three points. Let $G : \mathbb{R} \to \mathbb{R}$ obey the derivative estimates*

$$\left|\frac{d^j}{dx^j} G(x)\right| = O(n^{c_0}) \tag{8}$$

for $0 \leq j \leq 5$. Let $z_0 = E + \sqrt{-1}\eta_0$ be a complex number with $|E| \leq 2 - \delta$ for some fixed $\delta > 0$. Then

$$\mathbf{E}G(\log|\det(M_n - \sqrt{n}z_0)|) - \mathbf{E}G(\log|\det(M_n' - \sqrt{n}z_0)|) = O(n^{-c})$$

for some fixed $c > 0$, adopting the convention that $G(-\infty) = 0$.

If $E = 0$, then the requirement that the real and imaginary parts of the atom distribution are supported on at least three points can be dropped.

We prove this theorem in Section 4, following a preparation in Section 3. The requirements that M_n, M_n' be supported on at least three points, and that E lie in the bulk region $|E| < 2-\delta$ are artificial, due to the state of current literature on level repulsion estimates (see Proposition 14). It is likely that with further progress on those estimates that these hypotheses can be removed. The hypothesis that the atom distributions have independent real and imaginary parts is mostly for notational convenience and can also be removed with some additional effort. The hypothesis that the entries are bounded in magnitude by n^{c_0} is, strictly speaking, not satisfied for distributions such as the Gaussian distribution, but in practice we will be able to reduce to this case by a truncation argument.

By combining Theorem 5 with Theorem 4 we obtain

Corollary 6 (*Central Limit Theorem for Log-Determinant of Wigner Matrices*)**.** *Let M_n be a Wigner matrix whose atom distributions ζ_{ij} are independent of n, have real and imaginary parts that are independent and match GUE to fourth order, and obey Condition* **C1** *for some sufficiently large absolute constant C_0. Then*

$$\frac{\log|\det(M_n)| - \frac{1}{2}\log n! + \frac{1}{4}\log n}{\sqrt{\frac{1}{2}\log n}} \to N(0,1)_{\mathbb{R}}.$$

If M_n matches GOE instead of GUE, then one instead has

$$\frac{\log|\det(M_n)| - \frac{1}{2}\log n! + \frac{1}{4}\log n}{\sqrt{\log n}} \to N(0,1)_{\mathbb{R}}.$$

The deduction of this corollary from Theorems 5 and 4 is standard (closely analogous, for instance, to the proof of [35, Corollary 21], which establishes a similar central limit theorem for individual eigenvalues of a Wigner matrix) and is omitted. (Notice that in order for the atom variables of M_n match those of GUE to fourth order, these variables must have at least three points in their supports.)

1.1. Notation

Throughout this paper, n is a natural number parameter going off to infinity; in particular we will assume that $n \geq 100$ (so that $\log \log \log n$ is well-defined). A quantity is said to be *fixed* if it does not depend on n. We write $X = O(Y)$, $X \ll Y$, or $Y \gg X$ if one has $|X| \leq CY$ for some fixed C, and $X = o(Y)$ if one has $X/Y \to 0$ as $n \to \infty$. Absolute constants such as C_0 or c_0 are always understood to be fixed.

We say that an event E occurs with *high probability* if it occurs with probability $1 - O(n^{-c})$ for some fixed $c > 0$, and with *overwhelming probability* if it occurs with probability $1 - O(n^{-A})$ for all fixed $A > 0$.

2. The central limit theorem for GUE

We now prove Theorem 4. For notational reasons we shall take n to be even, but the argument below can easily be verified to also work with minor modifications when n is odd. We will use a method suggested to us by Rowan Killip (private communication), and loosely based on the arguments in [19].

We will work for most of this section with the GUE case, and discuss the changes in the numerology needed to address the GOE case at the end of the section.

The starting point is the following beautiful observation of Trotter [38]:

Proposition 7 (*Tridiagonal form of GUE [38]*)**.** *Let M_n' be the random tridiagonal real symmetric matrix*

$$M_n' = \begin{pmatrix} a_1 & b_1 & 0 & \cdots & 0 & 0 \\ b_1 & a_2 & b_2 & \cdots & 0 & 0 \\ 0 & b_2 & a_3 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & a_{n-1} & b_{n-1} \\ 0 & 0 & 0 & \cdots & b_{n-1} & a_n \end{pmatrix}$$

where the $a_1, \ldots, a_n, b_1, \ldots, b_{n-1}$ are jointly independent real random variables, with $a_1, \ldots, a_n \equiv N(0,1)_{\mathbb{R}}$ being standard real Gaussians, and each b_i having a complex χ-distribution:

$$b_i = \left(\sum_{j=1}^{i} |z_{i,j}|^2 \right)^{1/2}$$

where $z_{i,j} \equiv N(0,1)_{\mathbb{C}}$ are i.i.d. complex Gaussians.[5] *Let M_n be drawn from GUE. Then the joint eigenvalue distribution of M_n is identical to the joint eigenvalue distribution of M_n'.*

Proof. Let M_n be drawn from GUE. We can write

$$M_n = \begin{pmatrix} M_{n-1} & X_n \\ X_n^* & a_n \end{pmatrix}$$

where M_{n-1} is drawn from the $n-1 \times n-1$ GUE, $a_n \equiv N(0,1)_{\mathbb{R}}$, and $X_n \in \mathbb{C}^{n-1}$ is a random Gaussian vector with all entries i.i.d. with distribution $N(0,1)_{\mathbb{C}}$. Furthermore, M_{n-1}, X_n, a_n are jointly independent.

[5] In other words, the real and imaginary parts of $z_{i,j}$ are independent with distribution $N(0,1/2)_{\mathbb{R}}$.

We now apply the tridiagonal matrix algorithm. Let $b_{n-1} := |X_n|$, then b_n has the χ-distribution indicated in the proposition. We then conjugate M_n by a unitary matrix U that preserves the final basis vector e_n, and maps X_n to $b_{n-1}e_{n-1}$. Then we have

$$UM_nU^* = \begin{pmatrix} \tilde{M}_{n-1} & b_{n-1}e_{n-1} \\ b_{n-1}e_{n-1}^* & a_n \end{pmatrix}$$

where $\tilde{M}_{n-1}$ is conjugate to M_{n-1}. Now we make the crucial observation: because M_{n-1} is distributed according to GUE (which is a unitarily invariant ensemble), and U is a unitary matrix independent of M_{n-1}, $\tilde{M}_{n-1}$ is also distributed according to GUE, and remains independent of both b_{n-1} and a_n.

We continue this process, expanding UM_nU^* as

$$\begin{pmatrix} M_{n-2} & X_{n-1} & 0 \\ X_{n-1}^* & a_{n-1} & b_{n-1} \\ 0 & b_{n-1} & a_n \end{pmatrix}.$$

Applying a further unitary conjugation that fixes e_{n-1}, e_n but maps X_{n-1} to $b_{n-2}e_{n-2}$, we may replace X_{n-1} by $b_{n-2}e_{n-2}$ while transforming M_{n-2} to another GUE matrix $\tilde{M}_{n-2}$ independent of $a_n, b_{n-1}, a_{n-1}, b_{n-2}$. Iterating this process, we eventually obtain a coupling of M_n to M_n' by unitary conjugations, and the claim follows. □

In what follows, we are going to prove the limit law for the model M_n' and hence for M_n. Since b_i^2 has expectation i and variance[6] i, we can write it as

$$b_i^2 = i + \sqrt{i}c_i,$$

where c_i has mean 0 and variance 1.

By the properties of normal distribution and chi square distribution (or from concentration of measure inequalities), we have the following tail bound. There are constants $C_1, C_2 > 0$ such that for all $i \geq 1$ and $t \geq 0$, one has

$$\mathbf{P}(\max\{|a_i|, |c_i|\} \geq t) \leq C_1 \exp(-t^{C_2}). \tag{9}$$

Let M_i' denote the upper left $i \times i$ minor of M_n', and write $D_i := \det M_i'$. From cofactor expansion we have the recursion

$$D_i = a_i D_{i-1} - b_{i-1}^2 D_{i-2}$$

for all $i \geq 2$. To prove Theorem 4, we need to establish the law

$$\frac{\log|D_n| - \frac{1}{2}\log(n-1)! - \frac{1}{4}\log n}{\sqrt{\frac{1}{2}\log n}} \to N(0,1)_{\mathbb{R}}.$$

It will be convenient to skip the first few terms of this recursion. Let m be a sufficiently slowly growing integer-valued function of n (e.g. $m := \lfloor \log\log\log n \rfloor$ will suffice); we will only apply this recursion for $i \geq m$. Notice that $D_i \neq 0$ with probability one for all i.

6 Note that the more familiar *real* chi squared distribution χ_i^2 would have variance $2i$ here, but b_i^2 has the *complex* chi squared distribution which has variance i.

We then have

$$D_i = a_i D_{i-1} - (i-1+\sqrt{i-1}c_{i-1})D_{i-2}.$$

To mostly eliminate the $i-1$ factor, we introduce the normalised determinants

$$E_i := \frac{D_i}{\sqrt{i!}}$$

and conclude the recurrence

$$E_i = \frac{a_i}{\sqrt{i}} E_{i-1} - \left(\frac{\sqrt{i-1}}{\sqrt{i}} + \frac{c_{i-1}}{\sqrt{i}}\right) E_{i-2}.$$

By Taylor expansion we can rewrite this as

$$E_i = \frac{a_i}{\sqrt{i}} E_{i-1} - \left(1 + \frac{c_{i-1}}{\sqrt{i}} - \frac{1}{2i} + O\left(\frac{1}{i^2}\right)\right) E_{i-2}. \tag{10}$$

Our task is now to show that

$$\frac{\log|E_n| + \frac{1}{4}\log n}{\sqrt{\frac{1}{2}\log n}} \to N(0,1)_{\mathbb{R}}. \tag{11}$$

To deduce this central limit theorem from (10), we would like to write $\log|E_i|$ as a sum of martingale differences. But it is rather hard to do from the above recursive formula (10). We will need to perform an additional algebraic manipulation to obtain a more tractable formula involving the closely related quantity $F_j := E_{2j}^2 + E_{2j-1}^2$. In particular, we will establish

Proposition 8 (*Central Limit Theorem for $F_{n/2}$*)**.** *We have*

$$\frac{\log F_{n/2} + \frac{1}{2}\log n}{\sqrt{2\log n}} \to N(0,1)_{\mathbb{R}}. \tag{12}$$

We now prove this proposition. The idea is use Taylor expansions (which can be viewed as a discrete version of Ito's stochastic calculus) to approximate $\log F_{n/2} + \frac{1}{2}\log n$ as the sum $\sum_{j=1}^{n/2} \frac{1}{\sqrt{j}} h_j$ of martingale differences, to which the martingale central limit theorem may be applied.

We turn to the details. From (10) for $i = 2j, 2j-1$ we first observe the crude bound

$$F_j = O(Y_j^{O(1)} F_{j-1}), \tag{13}$$

where $Y_j := 1 + |a_{2j}| + |c_{2j-1}| + |a_{2j-1}| + |c_{2j-2}|$. Observe from (9) that

$$\mathbf{E} Y_j^l \ll 1 \tag{14}$$

for any fixed l.

Next, we apply (10) for $i = 2j, 2j-1$ and use Taylor expansion (using (13) to bound error terms of order $j^{-3/2}$ or better) to obtain

$$E_{2j} = \frac{a_{2j}}{\sqrt{2j}} E_{2j-1} - \left(1 + \frac{c_{2j-1}}{\sqrt{2j}} - \frac{1}{4j}\right) E_{2j-2} + O\left(\frac{Y_j^{O(1)}}{j^{3/2}} F_{j-1}^{1/2}\right)$$

$$E_{2j-1} = \left(\frac{a_{2j-1}}{\sqrt{2j}} + \frac{r_j^{[1]}}{j}\right) E_{2j-2} - \left(1 + \frac{c_{2j-2}}{\sqrt{2j}} - \frac{1}{4j} + \frac{r_j^{[2]}}{j}\right) E_{2j-3} + O\left(\frac{Y_j^{O(1)}}{j^{3/2}} F_{j-1}^{1/2}\right),$$

where $r_j^{[1]}, r_j^{[2]}$ are random variables bounded in magnitude by $O(Y_j^{O(1)})$ and with mean zero. (In fact, we can obtain a denominator of j^2 instead of $j^{3/2}$ in the error terms here, although this improved error term will not persist in our later analysis.) Substituting the second equation into the first (and again using (13) to handle all terms of order $j^{-3/2}$ or better), we also obtain

$$E_{2j} = -\left(1 + \frac{c_{2j-1}}{\sqrt{2j}} - \frac{1}{4j} + \frac{r_j^{[3]}}{j}\right) E_{2j-2} - \left(\frac{a_{2j}}{\sqrt{2j}} + \frac{r_j^{[4]}}{j}\right) E_{2j-3} + O\left(\frac{Y_j^{O(1)}}{j^{3/2}} F_{j-1}^{1/2}\right)$$

where $r_j^{[3]}, r_j^{[4]}$ obey the same sort of bounds as $r_j^{[1]}, r_j^{[2]}$. We may rewrite these estimates in matrix form as

$$\begin{pmatrix} E_{2j} \\ E_{2j-1} \end{pmatrix} = \left(-1 + \frac{1}{\sqrt{2j}} G_j + \frac{1}{4j} + \frac{1}{j} R_j\right) \begin{pmatrix} E_{2j-2} \\ E_{2j-3} \end{pmatrix} + O\left(\frac{Y_j^{O(1)}}{j^{3/2}} F_{j-1}^{1/2}\right) \tag{15}$$

where G_j is the near-Gaussian matrix

$$G_j := \begin{pmatrix} -c_{2j-1} & -a_{2j} \\ a_{2j-1} & -c_{2j-2} \end{pmatrix}, \tag{16}$$

and R_j is a random matrix depending on $j, a_{2j-1}, a_{2j}, c_{2j-2}, c_{2j-1}$ with mean zero and whose entries are bounded by $O(|Y_j|^{O(1)})$. (We remind the reader at this point that the implied constants in the $O()$ notation are independent of j.)

Using (15), we can express

$$F_j = \begin{pmatrix} E_{2j} & E_{2j-1} \end{pmatrix} \begin{pmatrix} E_{2j} \\ E_{2j-1} \end{pmatrix}$$

as

$$\begin{pmatrix} E_{2j-2} & E_{2j-3} \end{pmatrix} \left(-1 + \frac{1}{\sqrt{2j}} G_j + \frac{1}{4j} + \frac{1}{j} R_j\right)^* \times \left(-1 + \frac{1}{\sqrt{2j}} G_j + \frac{1}{4j} + \frac{1}{j} R_j\right) \begin{pmatrix} E_{2j-2} \\ E_{2j-3} \end{pmatrix} + O\left(\frac{Y_j^{O(1)}}{j^{3/2}} F_{j-1}\right).$$

We can collect some terms, splitting $G_j^* G_j$ as the sum of 2 and the mean zero random matrix $G_j^* G_j - 2$, and obtain the expansion[7]

$$F_j = \left(1 + \frac{\sqrt{2}}{\sqrt{j}} h_j + \frac{1}{2j} + \frac{1}{j} k_j + O\left(\frac{Y_j^{O(1)}}{j^{3/2}}\right)\right) F_{j-1} \tag{17}$$

where

$$h_j := \frac{1}{F_{j-1}} \begin{pmatrix} E_{2j-2} & E_{2j-3} \end{pmatrix} G_j \begin{pmatrix} E_{2j-2} \\ E_{2j-3} \end{pmatrix}$$
$$\frac{1}{F_{j-1}} \begin{pmatrix} E_{2j-2} & E_{2j-3} \end{pmatrix} G_j^* \begin{pmatrix} E_{2j-2} \\ E_{2j-3} \end{pmatrix}$$

and

$$k_j := \frac{1}{F_{j-1}} \begin{pmatrix} E_{2j-2} & E_{2j-3} \end{pmatrix} R'_j \begin{pmatrix} E_{2j-2} \\ E_{2j-3} \end{pmatrix}$$

and R'_j is a random matrix depending on $j, a_{2j-1}, a_{2j}, c_{2j-2}, c_{2j-1}$ with mean zero and entries bounded by $O(Y_j^{O(1)})$.

We can expand h_j as

$$\begin{aligned} &(-c_{2j-1}) \frac{E_{2j-2}^2}{E_{2j-2}^2 + E_{2j-3}^2} + (-c_{2j-2}) \frac{E_{2j-3}^2}{E_{2j-2}^2 + E_{2j-3}^2} \\ &\quad + (a_{2j-1} - a_{2j}) \frac{E_{2j-2} E_{2j-3}}{E_{2j-2}^2 + E_{2j-3}^2}. \end{aligned} \tag{18}$$

As the c_l, a_l are independent and all have mean 0 and variance 1, we conclude that for any fixed E_{2j-2} and E_{2j-3}, h_j also has mean zero and variance 1, thus

$$\mathbf{E}(h_j | \mathcal{E}_{j-1}) = 0; \qquad \mathbf{E}(h_j^2 | \mathcal{E}_{j-1}) = 1, \tag{19}$$

where $\mathcal{E}_l$ is the σ-algebra generated by the random variables $a_1, \ldots, a_{2l}$ and $c_1, \ldots, c_{2l-1}$ (or equivalently, by the entries of the minor M_{2l}). Similarly, for any fixed choice of E_{2j-2}, E_{2j-3}, k_j is a real random variable with mean zero, and thus

$$\mathbf{E}(k_j | \mathcal{E}_{j-1}) = 0. \tag{20}$$

Also, from construction, $h_j, k_j = O(Y_j^{O(1)})$.

Taking logarithms in (17), we obtain

$$\log F_j = \log F_{j-1} + \log\left(1 + \frac{1}{2j} + \frac{\sqrt{2}}{\sqrt{j}} h_j + \frac{1}{j} k_j + O\left(\frac{Y_j^{O(1)}}{j^{3/2}}\right)\right).$$

[7] The $\frac{1}{2j}$ term here arises from combining three contributions $-1 \times \frac{1}{4j} + \frac{1}{4j} \times (-1) + \frac{1}{\sqrt{2j}} \times \frac{1}{\sqrt{2j}} \times 2$.

By telescoping series, we may thus write

$$\log F_{n/2} = \log F_m + \sum_{j=m+1}^{n/2} \log\left(1 + x_j + O\left(\frac{Y_j^{O(1)}}{j^{3/2}}\right)\right)$$

where

$$x_j := \frac{1}{2j} + \frac{\sqrt{2}}{\sqrt{j}} h_j + \frac{1}{j} k_j. \tag{21}$$

For m sufficiently slowly growing in n, we clearly have

$$\log F_m = o(\sqrt{\log n})$$

with probability $1 - o(1)$, since F_m is almost surely finite with a law that depends only on m and not on n. To prove (12), it thus suffices to show that

$$\frac{\sum_{j=m+1}^{n/2} \log\left(1 + x_j + O\left(\frac{Y_j^{O(1)}}{j^{3/2}}\right)\right) + \frac{1}{2}\log n}{\sqrt{2\log n}} \to N(0,1)_{\mathbb{R}}. \tag{22}$$

The next step is to use Taylor expansion to approximate the logarithm to extract something that more closely resembles a martingale difference. Observe that $x_j = O(Y_j^{O(1)}/j^{1/2})$. From (14), we conclude that with probability $1 - O(j^{-100})$ (say), the expression $1 + x_j + O(\frac{Y_j^{O(1)}}{j^{3/2}})$ lies between $1/2$ and $3/2$ (say). From the union bound, we thus see that with probability

$$1 - \sum_{j=m+1}^{n/2} O(j^{-100}) = 1 - o(1),$$

one has

$$\log\left(1 + x_j + O\left(\frac{Y_j^{O(1)}}{j^{3/2}}\right)\right) = x_j - x_j^2/2 + O\left(\frac{Y_j^{O(1)}}{j^{3/2}}\right)$$

for all $m + 1 \le j \le n/2$. As h_j has variance 1, we can split h_j^2 as the sum of 1 and the mean zero random variable $h_j^2 - 1$. From (21) we may thus expand

$$x_j - x_j^2/2 = -\frac{1}{2j} + \frac{\sqrt{2}}{\sqrt{j}} h_j + \frac{1}{j} k_j' + O\left(\frac{Y_j^{O(1)}}{j^{3/2}}\right)$$

where k_j' is a random variable bounded by $O(Y_j^{O(1)})$ which has conditional mean zero:

$$\mathbf{E}(k_j' | \mathcal{E}_{j-1}) = 0.$$

Similarly, from (14) and the union bound again, the $O(Y_j^{O(1)}/j^{3/2})$ error terms are $O(1/j^{1.1})$ (say) with probability $1 - o(1)$, and thus we see that with probability $1 - o(1)$, we have

$$\sum_{j=m+1}^{n/2} \log\left(1 + x_j + \left(\frac{Y_j^{O(1)}}{j^{3/2}}\right)\right) = \sum_{j=m+1}^{n/2} \frac{\sqrt{2}}{\sqrt{j}} h_j + \frac{1}{j} k_j' - \frac{1}{2}\log n + O(1). \tag{23}$$

To prove (22), it thus suffices to show that

$$\frac{\sum_{j=m+1}^{n/2} \frac{\sqrt{2}}{\sqrt{j}} h_j + \frac{1}{j} k'_j}{\sqrt{2\log n}} \to N(0,1)_{\mathbb{R}}. \tag{24}$$

Observe that as each k'_j are martingale differences, which have variance $O(1)$ thanks to (14). As such, the expression $\sum_{j=m+1}^{n/2} \frac{1}{j} k'_j$ has variance $\sum_{j=m+1}^{n/2} O(1/j^2) = o(\log n)$ and can thus be discarded. If m is small enough, then the expression $\sum_{j=1}^{m} \frac{\sqrt{2}}{\sqrt{j}} h_j$ has variance $o(\log n)$ and can similarly be discarded. Thus it suffices to show that

$$\frac{\sum_{j=1}^{n/2} \frac{1}{\sqrt{j}} h_j}{\sqrt{\log n}} \to N(0,1)_{\mathbb{R}}. \tag{25}$$

In order to verify (25), we need to invoke the martingale central limit theorem:

Theorem 9 (*Martingale Central Limit Theorem*). *Assume that $T_1, \dots, T_n$ are martingale differences with respect to the nested σ-algebra $\mathcal{E}_0, \mathcal{E}_1, \dots, \mathcal{E}_{n-1}$. Let $v_n^2 := \sum_{i=1}^n \mathbf{E}(T_i^2 | \mathcal{E}_{i-1})$, and $s_m^2 := \sum_{i=1}^n \mathbf{E}(T_i^2)$. Assume that*

- *$v_n/s_n \to 1$ in probability;*
- *(Lindeberg condition) for every $\epsilon > 0$, $s_n^{-2} \sum_{i=0}^{n-1} \mathbf{E}(T_{i+1}^2 \mathbf{1}_{T_{i+1} \geq \epsilon s_n}) \to 0$ as $m \to \infty$.*

Then $\frac{\sum_{i=1}^n T_i}{s_n} \to N(0,1)_{\mathbb{R}}$.

Proof. See [5, Theorem 1]. □

We apply this theorem with $T_j := \frac{1}{\sqrt{j}} h_j$. From (19) one has $\mathbf{E}(T_{j+1} | \mathcal{E}_j) = 0$ and $\mathbf{E}(T_{j+1}^2 | \mathcal{E}_j) = \frac{1}{j}$, and hence also $\mathbf{E}(T_{j+1}^2) = \frac{1}{j}$. Thus $v_n = s_n = \log^{1/2} n + O(1)$; this gives the first hypothesis in Theorem 9.

Now we verify the Lindeberg condition. From (14) we have $\mathbf{E}(T_{i+1}^4 | \mathcal{E}_i) = \frac{1}{(i+1)^2} \mathbf{E}(h_{i+1}^4 | \mathcal{E}_i) \ll \frac{1}{i^2}$ and hence

$$\mathbf{E}(T_{i+1}^2 \mathbf{1}_{T_{i+1} \geq \epsilon s_n}) \leq \epsilon^{-2} s_n^{-2} \mathbf{E} T_{i+1}^4 \ll \frac{\epsilon^{-2}}{i^2 s_n};$$

since $s_n = \log^{1/2} n + O(1)$, the claim follows. This concludes the proof of Proposition 8.

Proposition 8 controls the magnitude of the vector $\binom{E_{2j}}{E_{2j-1}}$ when $j = n/2$. We will however be interested in the distribution of E_n, and so we must also obtain information about the *phase* of this vector also. To this end, we express this vector in polar coordinates as

$$\begin{pmatrix} E_{2j} \\ E_{2j-1} \end{pmatrix} = (-1)^j F_j^{1/2} \begin{pmatrix} \cos\theta_j \\ \sin\theta_j \end{pmatrix} \tag{26}$$

for some $\theta_j \in \mathbb{R}/2\pi\mathbf{Z}$, where we introduce the sign $(-1)^j$ to cancel the -1 factor in (15).

Proposition 10 (*Uniform Distribution of θ_n*). *One has*

$$\theta_n \to \mathbf{u}$$

as $n \to \infty$, where $\mathbf{u}$ is the uniform distribution on $\mathbb{R}/2\pi\mathbf{Z}$.

Proof. By the Weyl equidistribution criterion, it suffices to show that

$$\mathbf{E}e^{ik\theta_n} = o(1)$$

for every fixed non-zero integer k.

Fix k. Inserting the polar representation (26) into (15), we obtain the recursion

$$\begin{pmatrix} \cos\theta_j \\ \sin\theta_j \end{pmatrix} = C_j(1 + D_j) \begin{pmatrix} \cos\theta_{j-1} \\ \sin\theta_{j-1} \end{pmatrix} \tag{27}$$

where D_j is the matrix

$$D_j := -\frac{1}{\sqrt{2j}} G_j + \frac{1}{j} R''_j + O\left(\frac{Y_j^{O(1)}}{j^{3/2}}\right), \tag{28}$$

where R''_j is a matrix obeying the same properties as R_j or R'_j, and C_j is a non-zero scalar whose exact value is not important for us.

We extract the components e_j, f_j of D_j in the orthonormal basis formed by the two vectors $(\cos\theta_{j-1}, \sin\theta_{j-1})$ and $(-\sin\theta_{j-1}, \cos\theta_{j-1})$, thus

$$e_j := \begin{pmatrix} -\sin\theta_{j-1}, & \cos\theta_{j-1} \end{pmatrix} D_j \begin{pmatrix} \cos\theta_{j-1} \\ \sin\theta_{j-1} \end{pmatrix}$$

and

$$f_j := \begin{pmatrix} \cos\theta_{j-1}, & \sin\theta_{j-1} \end{pmatrix} D_j \begin{pmatrix} \cos\theta_{j-1} \\ \sin\theta_{j-1} \end{pmatrix}.$$

From (27) we thus have

$$\frac{1}{C_j}(\cos\theta_j, \sin\theta_j) = (1 + f_j)(\cos\theta_{j-1}, \sin\theta_{j-1}) + e_j(-\sin\theta_{j-1}, \cos\theta_{j-1}),$$

and so we have a right-angled triangle with base $1+f_j$, height e_j, and angle $\theta_j - \theta_{j-1}$. Elementary trigonometry then gives

$$\tan(\theta_j - \theta_{j-1}) = \frac{e_j}{1 + f_j}.$$

By (14), we see that with probability $1 - O(j^{-100})$ (say), we have $e_j, f_j = O(j^{-0.49})$ (say). Using the Taylor expansion of arc tan and $1/(1 + x)$ we obtain

$$\theta_j - \theta_{j-1} = e_j - e_j f_j + O(j^{-1.47})$$

and thus

$$e^{ik\theta_j} = e^{ik\theta_{j-1}}\left(1 + ike_j - ike_j f_j - \frac{k^2}{2}e_j^2 + O(j^{-1.47})\right)$$

with probability $1 - O(j^{-100})$. Hence

$$\mathbf{E}e^{ik\theta_j} = \mathbf{E}e^{ik\theta_{j-1}}\left(1 + ike_j - ike_j f_j - \frac{k^2}{2}e_j^2\right) + O(j^{-1.47}).$$

Now from (28), (16), (14) we see that after conditioning on θ_{j-1}, e_j has mean $O(j^{-1.47})$ and variance $\frac{1}{2j} + O(j^{-1.47})$, and that $e_j f_j$ has mean $O(j^{-1.47})$. We conclude that

$$\mathbf{E}e^{ik\theta_j} = \mathbf{E}e^{ik\theta_{j-1}}\left(1 - \frac{k^2}{4j}\right) + O(j^{-1.47})$$

for any $1 \le j \le n$. Telescoping this, we see that

$$|\mathbf{E}e^{ik\theta_n}| \ll \left(\frac{m}{n}\right)^{k^2/4} |\mathbf{E}e^{ik\theta_m}| + O(m^{-0.47})$$

for any $1 \le m \le n$. Bounding $|\mathbf{E}e^{ik\theta_m}|$ by 1 and choosing m to be a slowly growing function of n, we obtain the claim. $\square$

From the above proposition, we see in particular that

$$\frac{1}{\log n} \le |\cos\theta_{n/2}| \le 1$$

(say) with probability $1 - o(1)$. Since $E_n = (-1)^{n/2} F_{n/2}^{1/2} \cos\theta_{n/2}$, we thus have

$$\log|E_n| = \frac{1}{2}\log F_{n/2} + O(\log\log n)$$

with probability $1 - o(1)$; combining this with Proposition 8, we see that

$$\frac{\log|E_n|^2 + \frac{1}{2}\log n}{\sqrt{2\log n}} \to N(0,1)_{\mathbb{R}}.$$

The claim (11) then follows.

2.1. The GOE case

We now discuss the changes to the above argument needed to address the GOE case. The analogue of Proposition 7 is easily established, but with the changes that the a_j now have the distribution of $N(0,2)_{\mathbb{R}}$ instead of $N(0,1)_{\mathbb{R}}$, and the b_j now have a real χ-distribution instead of a complex one (thus the $z_{i,j}$ are now distributed according to $N(0,1)_{\mathbb{R}}$ instead of $N(0,1)_{\mathbb{C}}$). The effect of this is to make the random variables a_j, c_j in the above analysis have variance 2 instead of 1 (but they still have mean zero). As a consequence G^*G now has mean 4 rather than mean 2, which means that the $\frac{1}{2j}$ term in (17) becomes $\frac{3}{2j}$. On the other hand, the random variables h_j now have variance 2 instead of 1. These two changes cancel each other out to some extent, and in particular the assertion (23) remains unchanged. Finally, when applying the martingale central limit theorem, the variances v_n^2, s_n^2 are now $2\log n + O(1)$ rather than $\log n + O(1)$, again thanks to the increased variance of h_j. The remainder of the argument goes through with the obvious changes.

3. Resolvent swapping: a deterministic analysis

In this section we study the stability of Hermitian matrices with respect to perturbation in just one or two entries. To formalise this we will need some definitions.

We will need a number of matrix norms. Let $A = (a_{ij})_{1\le i,j\le n}$ be a matrix, and let $1 \le p, q \le \infty$ be exponents. We use $\|A\|_{(q,p)}$ to denote the $\ell^p \to \ell^q$ operator norm, i.e. the best constant in the inequality

$$\|Ax\|_{\ell^q} \le \|A\|_{(q,p)}\|x\|_{\ell^p}.$$

Thus for instance $\|A\|_{(2,2)}$ is the usual operator norm. We also observe the identities

$$\|A\|_{(\infty,1)} = \sup_{1\le i,j\le n} |a_{ij}|$$

and

$$\|A\|_{(\infty,2)} := \sup_{1\le i\le n} \left(\sum_{j=1}^{n} |a_{ij}|^2\right)^{1/2}.$$

In particular one has the identity

$$\|A\|_{(\infty,2)} = \|AA^*\|_{(\infty,1)}^{1/2}. \tag{29}$$

By duality one has

$$\|A\|_{(q,p)} = \|A^*\|_{(p',q')}, \tag{30}$$

where $1/p + 1/p' = 1/q + 1/q' = 1$.

We observe the trivial inequality

$$\|AB\|_{(r,p)} \le \|A\|_{(r,q)}\|B\|_{(q,p)} \tag{31}$$

for any A, B and $1 \le p, q, r \le \infty$.

Next, we need the notion of an elementary matrix.

Definition 11 (*Elementary Matrix*)**.** An *elementary matrix* is a matrix which has one of the following forms

$$V = e_a e_a^*,\ e_a e_b^* + e_b e_a^*,\ \sqrt{-1}e_a e_b^* - \sqrt{-1}e_b e_a^* \tag{32}$$

with $1 \le a, b \le n$ distinct, where $e_1, \dots, e_n$ is the standard basis of $\mathbb{C}^n$.

Observe that

$$\|V\|_{(q,p)} \ll 1 \tag{33}$$

and

$$|\mathrm{trace}(AV)| = |\mathrm{trace}(VA)| = O(\|A\|_{(q,p)}) \tag{34}$$

for all $1 \le p, q \le \infty$ and all $n \times n$ matrices A.

Let M_0 be a Hermitian matrix, let $z = E + i\eta$ be a complex number, and let V be an elementary matrix. We then introduce, for each $t \in \mathbb{R}$, the Hermitian matrices

$$M_t := M_0 + \frac{1}{\sqrt{n}} tV,$$

the resolvent

$$R_t = R_t(E + i\eta) := (M_t - E - i\eta)^{-1} \tag{35}$$

and the Stieltjes transform

$$s_t := s_t(E + i\eta) := \frac{1}{n}\mathrm{trace} R_t(E + i\eta)$$

and study how R_t and s_t depend on t.

We have the fundamental *resolvent identity*

$$R_t = R_0 - \frac{t}{\sqrt{n}} R_0 V R_t$$

which upon iteration leads to

$$R_t = R_0 + \sum_{j=1}^{k}\left(-\frac{t}{\sqrt{n}}\right)^j (R_0 V)^j R_0 + \left(-\frac{t}{\sqrt{n}}\right)^{k+1} (R_0 V)^{k+1} R_t. \tag{36}$$

Under a mild hypothesis, we also have the infinite limit of (36):

Lemma 12 (*Neumann Series*)**.** *Let M_0 be a Hermitian $n \times n$ matrix, let $E \in \mathbb{R}, \eta > 0$, and $t \in \mathbb{R}$, and let V be an elementary matrix. Suppose one has*

$$|t|\|R_0\|_{(\infty,1)} = o(\sqrt{n}). \tag{37}$$

Then one has the Neumann series formula

$$R_t = R_0 + \sum_{j=1}^{\infty}\left(-\frac{t}{\sqrt{n}}\right)^j (R_0 V)^j R_0 \tag{38}$$

with the right-hand side being absolutely convergent, where R_t is defined by (35). Furthermore, for any $1 \le p \le \infty$ one has

$$\|R_t\|_{(\infty,p)} \le (1 + o(1))\|R_0\|_{(\infty,p)}. \tag{39}$$

In practice, we will have $t = n^{O(c_0)}$ (from a decay hypothesis on the atom distribution) and $\|R_0\|_{(\infty,1)} = n^{O(c_0)}$ (from eigenvector delocalisation and a level repulsion hypothesis), where $c_0 > 0$ is a small constant, so (37) is quite a mild condition. We also remark that by replacing M_0 and t with $M_0 + tV$ and $-t$ respectively, one can swap the roles of R_0 and R_t in the above lemma without difficulty.

Proof. We rewrite (36) as

$$\left(1 - \left(-\frac{t}{\sqrt{n}}\right)^{k+1} (R_0 V)^{k+1}\right) R_t = R_0 + \sum_{j=1}^{k}\left(-\frac{t}{\sqrt{n}}\right)^j (R_0 V)^j R_0 \tag{40}$$

for all $k \ge 0$. Sending $k \to \infty$ we will be able to conclude (38) (in a conditionally convergent sense, at least) once we show that $\left(-\frac{t}{\sqrt{n}}\right)^{k+1} (R_0 V)^{k+1}$ converges to zero (in, say, $\|A\|_{(\infty,1)}$ norm) as $k \to \infty$. But from (33), (31) we have

$$\left\|\left(-\frac{t}{\sqrt{n}}\right)^{k+1} (R_0 V)^{k+1}\right\|_{(\infty,1)} \le O\left(\frac{|t|}{\sqrt{n}}\right)^{k+1} \|R_0\|_{\infty}^{k+1}.$$

From (37), this decays exponentially in k, and this gives (38) (and also demonstrates that the series is absolutely convergent).

Taking (∞, p) norms of (38), one has

$$\|R_t\|_{(\infty,p)} \le \|R_0\|_{(\infty,p)} + \sum_{j=1}^{\infty} \left(\frac{|t|}{\sqrt{n}}\right)^j \|(R_0V)^j R_0\|_{(\infty,p)}.$$

But from (33), (31) one has

$$\|(R_0V)^j R_0\|_{(\infty,p)} \le (\|R_0\|_{\infty,1} \|V\|_{1,\infty}\ \|)^j \|R_0\|_{\infty,p} = O(\|R_0\|_{\infty,1})^j \|R_0\|_{(\infty,p)}$$

and the claim (39) follows from (37). □

We now can describe the dependence of s_t on t:

Proposition 13 (*Taylor Expansion of s_t*)**.** *Let the notation be as above, and suppose that (37) holds. Let $k \ge 0$ be fixed. Then one has*

$$s_t = s_0 + \sum_{j=1}^{k} n^{-j/2} c_j t^j + O\left(n^{-(k+1)/2}|t|^{k+1}\|R_0\|_{(\infty,1)}^{k+1} \min\left(\|R_0\|_{(\infty,1)}, \frac{1}{n\eta}\right)\right) \tag{41}$$

where the coefficients c_j are independent of t and obey the bounds

$$c_j = O\left(\|R_0\|_{(\infty,1)}^{j} \min\left\{\|R_0\|_{(\infty,1)}, \frac{1}{n\eta}\right\}\right) \tag{42}$$

for all $1 \le j \le k$.

Proof. If we take normalised traces of (36), we obtain

$$s_t = s_0 + \sum_{j=1}^{k} n^{-j/2} c_j t^j + n^{-(k+1)/2} t^{k+1} r_t^{(k)}$$

where c_j are the coefficients

$$c_j := (-1)^j \frac{1}{n} \operatorname{trace}((R_0V)^j R_0)$$

and $r_t^{(k)}$ is the error term

$$r_t^{(k)} := (-1)^{k+1} \frac{1}{n} \operatorname{trace}((R_0V)^{k+1} R_t).$$

To estimate these coefficients, we use the cyclic property of trace to rearrange

$$\operatorname{trace}((R_0V)^j R_0) = \operatorname{trace}(V(R_0V)^{j-1} R_0^2)$$

and thus by (34), (31), (33) we have

$$\begin{aligned}
|c_j| &\ll \frac{1}{n}\|(R_0V)^{j-1}R_0^2\|_{(\infty,1)} \\
&\ll \frac{1}{n}\|R_0V\|_{(\infty,\infty)}^{j-1}\|R_0^2\|_{(\infty,1)} \\
&\ll \frac{1}{n}(\|R_0\|_{(\infty,1)}\|V\|_{(1,\infty)})^{j-1}\|R_0^2\|_{(\infty,1)} \\
&\ll \frac{1}{n}(\|R_0\|_{(\infty,1)}^{j-1}\|R_0^2\|_{(\infty,1)}).
\end{aligned}$$

We can bound this in one of two ways. Firstly, by (31) we have the crude inequality

$$\|R_0^2\|_{(\infty,1)} \le \|R_0\|_{(\infty,\infty)}\|R_0\|_{(\infty,1)} \le n\|R_0\|_{(\infty,1)}^2$$

(coming from the bound $\|x\|_{\ell^1} \le n\|x\|_{\ell^\infty}$), leading to the bound

$$c_j = O(\|R_0\|_{(\infty,1)}^{j+1}).$$

Alternatively, we can use (31), (30), (29) to bound

$$\begin{aligned}\|R_0^2\|_{(\infty,1)} &\le \|R_0\|_{(\infty,2)}\|R_0\|_{(2,1)}\\ &\le \|R_0\|_{(\infty,2)}\|R_0^*\|_{(\infty,2)}\\ &\le \|R_0R_0^*\|_{(\infty,1)}^{1/2}\|R_0^*R_0\|_{(\infty,1)}^{1/2}.\end{aligned}$$

But from the definition (35) of the resolvent, one has the identity

$$R_0R_0^* = R_0^*R_0 = \frac{R_0 - R_0^*}{2i\eta}$$

and thus from the triangle inequality and (30)

$$\|R_0^2\|_{(\infty,1)} \le \frac{1}{\eta}\|R_0\|_{(\infty,1)}.$$

This gives

$$c_j = O\left(\frac{1}{n\eta}\|R_0\|_{(\infty,1)}^j\right).$$

Combining the two bounds on c_j we have (42). A similar argument can be used to bound $r_t^{(k)}$ (using (39) to replace R_t by R_0 at some stage of the argument), so that

$$r_t^{(k)} = O\left(\|R_0\|_{(\infty,1)}^{k+1}\min\left\{\|R_0\|_{(\infty,1)}, \frac{1}{n\eta}\right\}\right). \tag{43}$$

The claim (41) follows. □

4. Proof of Theorem 5

In this section we prove Theorem 5. Let M_n, M_n' be as in that theorem, with c_0 sufficiently small to be chosen later. Call a statistic $S(M)$ that can depend on a matrix M *highly insensitive* if one has

$$|S(M_n) - S(M_n')| = O(n^{-c})$$

for some fixed $c > 0$. Thus our task is to show that $\mathbf{E}G(\log|\det(M_n - \sqrt{n}z_0)|)$ is highly insensitive for all z_0 and all G obeying (8). By dividing G by n^{c_0} (and reducing the size of c_0 if necessary) we may improve (8) to the estimates

$$\left|\frac{d^j}{dx^j}G(x)\right| \le 1 \tag{44}$$

for all $x \in \mathbb{R}$ and $0 \le j \le 5$.

By truncating the atom distributions (and re-adjusting to keep them at mean zero and unit variance) and using Condition **C1**, we may assume without loss of generality that we have the uniform upper bound

$$|\xi| \ll n^{c_0} \tag{45}$$

on the atom distribution (see [2, Chapter 2] or [26, Appendix A] for more details on the truncation technique).

Set $W_n := \frac{1}{\sqrt{n}} M_n$ (and $W_n' := \frac{1}{\sqrt{n}} M_n'$). Then

$$\log|\det(M_n - \sqrt{n} z_0)| = \frac{1}{2} n \log n + \log|\det(W_n - z_0)|.$$

By translating G by $\frac{1}{2} n \log n$ (which does not affect the bounds (44)), it thus suffices to show that $\mathbf{E}G(\log|\det(W_n - z_0)|)$ is highly insensitive.

Write $z_0 = E + \sqrt{-1}\eta_0$. By conjugation symmetry we may take $\eta_0 \geq 0$. We first dispose of the easy case when $\eta_0 \geq n^{100}$. In this case we have

$$\log|\det(W_n - z_0)| = n \log|z_0| + \log|\det(1 - z_0^{-1} W_n)| = n \log|z_0| + O(n^{-50})$$

(say), thanks to (45). The claim then follows easily in this case from (44).

We now restrict to the main case $0 \leq \eta_0 < n^{100}$. From the fundamental theorem of calculus one has

$$\log|\lambda - z_0| = \log|\lambda - E - \sqrt{-1} n^{100}| - \mathrm{Im} \int_{\eta_0}^{n^{100}} \frac{d\eta}{\lambda - E - \sqrt{-1}\eta}$$

and hence

$$\begin{aligned} \log|\det(W_n - z_0)| &= \log|\det(W_n - E - \sqrt{-1} n^{100})| \\ &\quad - n \mathrm{Im} \int_{\eta_0}^{n^{100}} s(E + \sqrt{-1}\eta)\, d\eta \end{aligned} \tag{46}$$

where

$$s(z) = s_{W_n}(z) = \frac{1}{n} \mathrm{trace}(W_n - z)^{-1}$$

is the Stieltjes transform of W_n.

The previous analysis and (46) then gives

$$\log|\det(W_n - z_0)| = n \log n^{100} + O(n^{-50}) - n \mathrm{Im} \int_{\eta_0}^{n^{100}} s(E + \sqrt{-1}\eta)\, d\eta.$$

By translating (and reflecting) G once more, it thus suffices to show that the quantity

$$\mathbf{E}G\left(n \mathrm{Im} \int_{\eta_0}^{n^{100}} s(E + \sqrt{-1}\eta)\, d\eta \right)$$

is highly insensitive.

We next need the following proposition. Let $\lambda_1(W_n) \geq \cdots \geq \lambda_n(W_n)$ denote the eigenvalues of W_n (counting multiplicity), and let $u_1(W_n), \ldots, u_n(W_n)$ be an associated orthonormal basis of eigenvectors.

Proposition 14 (*Non-Concentration*). *With high probability, one has*

$$\min_{1\le i\le n} |\lambda_i(W_n) - E| \ge n^{-1-c_0}$$

and with overwhelming probability one has

$$N_I = O(n|I|)$$

whenever I is an interval of length $|I| \ge n^{-1+Ac_0}$ for a sufficiently large constant $A > 0$. Also, with overwhelming probability one has

$$\sup_{1\le i\le n} \|u_i(W_n)\|_{\ell^\infty} \le n^{-1/2+O(c_0)}.$$

Proof. The second claim follows from [35, Proposition 66] and the third claim follows from [35, Proposition 62], so we turn to the first claim.

The results in [25] only give a lower bound of n^{-C} for some fixed C, which is not quite enough for our purposes. On the other hand, if the atom distribution is sufficiently smooth, the claim follows from existing level repulsion estimates such as [24] or [9], which are valid in the bulk region $|E| \le 2 - \delta$. To extend to the case when the real and imaginary parts of the atom distribution are supported on at least three points, one can use the Four Moment Theorem (see [35]) and a moment matching argument (see e.g. the proof of [35, Corollary 24]). We remark that these are the only places in which we use the hypotheses that $|E| \le 2 - \delta$ and that the real and imaginary parts of the distribution are supported on at least three points. It is likely that by improving the results in the above cited literature, one can remove these hypotheses.[8] □

As a consequence of Proposition 14, we obtain a upper bound on the (imaginary part of the) Stieltjes transform:

Corollary 15. *For a sufficiently large constant $A_0 > 0$ (independent of c_0), one has*

$$\operatorname{Im} s(E + \sqrt{-1}n^{-1-2A_0c_0}) \le n^{-A_0c_0}/2 \tag{47}$$

with high probability.

Proof. The left-hand side of (47) can be written as

$$n^{-2-2A_0c_0} \sum_{i=1}^{n} \frac{1}{n^{-2-4A_0c_0} + (\lambda_i(W_n) - E)^2}.$$

By Proposition 14, we assume with high probability that there are at most $O(n^{Ac_0})$ eigenvalues $\lambda_i(W_n)$ that are within n^{-1+Ac_0} of E, but that all such eigenvalues are at least n^{-1-c_0} away from E. The total contribution of these eigenvalues to the above expression is then at most $O(n^{(-2A_0+A+2)c_0})$. Similarly, by using Proposition 14 and dyadic decomposition of the spectrum around E, we see that the contribution of the eigenvalues that are further than n^{-1+Ac_0} away are $O(n^{(-2A_0-A)c_0})$ with overwhelming probability. Combining these bounds we obtain the claim if A_0 is large enough. □

[8] For instance, the results in [40] do not need the support hypothesis, but require the energy E to be zero and the imaginary part to vanish. It may however be possible to remove these hypotheses from the results in [40], which could lead to an improvement of the proposition here.

Let $\chi : \mathbb{R} \to \mathbb{R}$ be a smooth cutoff to the region $|x| \leq n^{-A_0 c_0}$ that equals 1 for $|x| \leq n^{-A_0 c_0}/2$. From the above corollary, $\chi(\mathrm{Im} s(E + \sqrt{-1} n^{-1-2A_0 c_0}))$ is equal to 1 with high probability. Thus it suffices to show that

$$\mathbf{E} G\left(n \mathrm{Im} \int_{\eta_0}^{n^{100}} s(E + \sqrt{-1}\eta)\, d\eta\right) \chi\left(\mathrm{Im} s(E + \sqrt{-1} n^{-1-2A_0 c_0})\right) \tag{48}$$

is highly insensitive.

We now view M'_n as obtained from M_n by n^2 swapping operations, each of which either replaces a diagonal entry of M_n with the corresponding entry of M'_n, or replaces the real or imaginary part of an off-diagonal entry of M_n (and its adjoint) with the corresponding entries of M'_n (leaving the other part of that entry unchanged). Of these n^2 swapping operations, n of them will involve a diagonal entry, and the other $n^2 - n$ will involve an off-diagonal entry. It will suffice to show that each swapping operation only affects (48) by $O(n^{-2-c})$ in the off-diagonal case and $O(n^{-1-c})$ in the diagonal case for some fixed $c > 0$. In fact we will obtain a bound of the form $O(n^{-5/2+O(c_0)})$ (where the implied constant may depend on A, A_0) in the off-diagonal case and $O(n^{-3/2+O(c_0)})$ in the diagonal case, which suffices for c_0 small enough.

Let $M_n^{(0)}, M_n^{(1)}$ be two adjacent matrices in this swapping process, and let $W_n^{(0)}, W_n^{(1)}$ be the associated normalised matrices. Then we can write

$$W_n^{(0)} = W_0 + \frac{1}{\sqrt{n}} \xi^{(0)} V$$
$$W_n^{(1)} = W_0 + \frac{1}{\sqrt{n}} \xi^{(1)} V$$

where V is an elementary matrix, $\xi^{(0)}, \xi^{(1)}$ are real random variables matching to fourth order and bounded in magnitude by $O(n^{O(c_0)})$, and W_0 is a random matrix independent of both $\xi^{(0)}$ and $\xi^{(1)}$. We can then write (48) for $W_n^{(0)}$ using the notation of the preceding section as

$$\mathbf{E} G\left(n \mathrm{Im} \int_{\eta_0}^{n^{100}} s_{\xi^{(0)}}(E + \sqrt{-1}\eta)\, d\eta\right) \chi(\mathrm{Im} s_{\xi^{(0)}}(E + \sqrt{-1} n^{-1-2A_0 c_0}))$$

and we wish to show that this expression only changes by $O(n^{-5/2+O(c_0)})$ when $\xi^{(0)}$ is replaced by $\xi^{(1)}$ in the off-diagonal case, or $O(n^{-3/2+O(c_0)})$ in the diagonal case.

We now bound the resolvent:

Lemma 16 (*Resolvent Bound*). *If $\chi(s_{\xi^{(0)}}(E + i n^{-1-2A_0 c_0}))$ is non-vanishing, then with overwhelming probability*

$$\sup_{\eta > 0} \|R_{\xi^{(0)}}(E + \sqrt{-1}\eta)\|_{(\infty,1)} \ll n^{O(c_0)}$$

and

$$\sup_{\eta > 0} \|R_0(E + i\eta)\|_{(\infty,1)} \ll n^{O(c_0)}. \tag{49}$$

Proof. From spectral decomposition one has

$$\|R_\xi(E + \sqrt{-1}\eta)\|_{(\infty,1)} \leq \sum_{i=1}^{n} \frac{\|u_j(W_n^{(0)})\|_{\ell^\infty}^2}{|\lambda_i(W_n) - E - \sqrt{-1}\eta|}.$$

Applying the last statement[9] in Proposition 14, we conclude with overwhelming probability that

$$\|R_\xi(E+\sqrt{-1}\eta)\|_{(\infty,1)} \le n^{-1+O(c_0)} \sum_{i=1}^{n} \frac{1}{|\lambda_i(W_n)-E|}.$$

Arguing as in Corollary 15, one see that if $\chi(s_\xi(E+\sqrt{-1}n^{-1-2A_0c_0}))$ is non-vanishing, then with overwhelming probability

$$\sum_{i=1}^{n} \frac{1}{|\lambda_i(W_n)-E|} = O(n^{1+O(c_0)})$$

and the first claim follows. The second claim then follows from Lemma 12 (swapping the roles of 0 and $\xi^{(0)}$). $\square$

We now condition to the event that (49) holds. To begin with, let us assume we are in the off-diagonal case. Then by Proposition 13 we have

$$s_{\xi^{(i)}}(E+\sqrt{-1}\eta) = s_0(E+\sqrt{-1}\eta) + \sum_{j=1}^{4} (\xi^{(i)})^j n^{-j/2} c_j(\eta) + O(n^{-5/2+O(c_0)}) \min\left\{1, \frac{1}{n\eta}\right\}$$

for $i=0,1$, where the coefficients c_j enjoy the bounds

$$c_j = O\left(n^{O(c_0)} \min\left\{1, \frac{1}{n\eta}\right\}\right).$$

From this and Taylor expansion above we see that the expression

$$G\left(n\operatorname{Im}\int_0^{n^{100}} s_{\xi^{(i)}}(E+\sqrt{-1}\eta)\,d\eta\right)\chi\left(\operatorname{Im} s_\xi(E+\sqrt{-1}n^{-1-2c_0})\right)$$

is equal to a polynomial of degree at most 4 in $\xi^{(i)}$ with coefficients independent of $\xi^{(i)}$, plus an error of $O(n^{-5/2+O(c_0)})$. Taking the expectation and using the four moment assumption, we obtain that the difference between the expectations of G with respect to $\xi^{(0)}$ and $\xi^{(1)}$ is $O(n^{-5/2+O(c_0)})$, as desired.

In the diagonal case, one argues similarly, except that one only is assuming two matching moments, and so one should only Taylor expand to second order rather than fourth order. This concludes the proof of Theorem 5.

Acknowledgments

We thank Brad Rodgers and Zhigang Bao for references, Peter Eichelsbacher, Xiuyuan Cheng, and the anonymous referee for corrections, and Rowan Killip for suggesting and explaining the tridiagonal method. T. Tao is supported by a grant from the MacArthur Foundation, by NSF grant DMS-0649473, and by the NSF Waterman award. V. Vu is supported by research grants DMS-0901216 and AFOSAR-FA-9550-09-1-0167.

[9] To be precise, we need to apply this statement for $W_n^{(0)}$, but the proof for this matrix is the same as for W_n; see [35].

Appendix. Moment calculations

In this Appendix we establish Theorem 3. Our main tool is the Leibniz expansion

$$\det M_n = \sum_{\sigma \in S_n} I_\sigma,$$

where for each permutation $\sigma : \{1, \ldots, n\} \to \{1, \ldots, n\}$, I_σ is the random variable

$$I_\sigma := \operatorname{sgn}(\sigma) \prod_{i=1}^{n} \zeta_{i\sigma(i)}.$$

We begin with the first moment computation. Clearly

$$\mathbf{E} \det M_n = \sum_{\sigma \in S_n} \mathbf{E} I_\sigma.$$

Because all the ζ_{ij} have mean zero and are jointly independent on the upper triangular region $1 \le i \le j$, we see that $\mathbf{E} I_\sigma$ vanishes unless σ consists entirely of 2-cycles (i.e. is a perfect matching), in which case $\mathbf{E} I_\sigma = 1$. Thus, $\mathbf{E} \det M_n$ is the number of perfect matchings on $\{1, \ldots, n\}$, which is easily seen to be zero when n is odd and $\frac{n!}{(n/2)! 2^{n/2}}$ when n is even.

Now we turn to the second moment computation for GOE, thus we seek the bounds

$$n^{3/2} n! \ll \mathbf{E} |\det M_n|^2 \ll n^{3/2} n! \tag{A.1}$$

We may of course assume that n is large.

From the Leibniz expansion we have

$$\mathbf{E} |\det M_n|^2 = \sum_{\sigma, \rho \in S_n} \mathbf{E} I_\sigma \overline{I_\rho}. \tag{A.2}$$

Actually, as GOE has real coefficients, we can omit the complex conjugate over the I_ρ term.

Now we investigate the expressions $\mathbf{E} I_\sigma I_\rho$. This expression can be estimated using the cycle decomposition of σ and ρ. It is not difficult to see that this expression will be zero unless the following conditions are satisfied:

- If γ is a cycle in σ of length other than two, then either γ or its reversal γ^{-1} is a cycle in ρ, and conversely.
- The support of the 2-cycles in σ equals the support of the 2-cycles in ρ.

Furthermore, if the above conditions are satisfied, then $\mathbf{E} I_\sigma \overline{I_\rho}$ is equal to $2^{C_1} 3^c$, where C_1 is the number of 1-cycles of σ (or of ρ), and c is the number of 2-cycles that are common to both σ and ρ. This comes from the fact that the diagonal entries of GOE have variance 2, while the off-diagonal entries have a fourth moment of 3.

Some elementary combinatorics shows that for a given permutation σ, the number of permutations ρ obeying the above conditions is equal to

$$\frac{2C_2!}{C_2! 2^{C_2}} \prod_{k \ge 3} 2^{C_k},$$

where C_k is the number of k-cycles of σ. (To be more precise, we would have to write $C_k(\sigma)$, but as there is little chance of misunderstanding, we prefer using just C_k to simplify the presentation.)

Thus (A.2) is thus lower bounded by

$$\sum_{\sigma\in S_n} 2^{C_1}\frac{2C_2!}{C_2!2^{C_2}}\prod_{k\geq 3}2^{C_k}. \tag{A.3}$$

In the converse direction, for fixed $0\leq c\leq C_2$, the number of ρ obeying the above conditions and with exactly c 2-cycles in common is bounded by

$$\binom{C_2}{c}\frac{2(C_2-c)!}{(C_2-c)!2^{C_2-c}}\prod_{k\geq 3}2^{C_k},$$

and so (A.2) is upper bounded by

$$\sum_{\sigma\in S_n} 2^{C_1}\sum_{c=0}^{C_2}3^c\binom{C_2}{c}\frac{2(C_2-c)!}{(C_2-c)!2^{C_2-c}}\prod_{k\geq 3}2^{C_k}.$$

Let us first estimate the upper bound. Observe from Stirling's formula that

$$\frac{2(C_2-c)!}{(C_2-c)!2^{C_2-c}}\ll (C_2-c)!2^{C_2-c}/\sqrt{C_2-c+1}.$$

This and an elementary calculation show

$$\sum_{c=0}^{C_2}3^c\binom{C_2}{c}\frac{2(C_2-c)!}{(C_2-c)!2^{C_2-c}}\ll C_2!2^{C_2}/\sqrt{C_2+1}.$$

So we may upper bound (A.2) by

$$O\left(\sum_{\sigma\in S_n}\frac{C_2!2^{C_2}}{\sqrt{C_2+1}}\prod_{k\neq 2}2^{C_k}\right).$$

Using the fact that $\sum_{m=1}^{n/2}\frac{1}{\sqrt{m}}=O(\sqrt{n})$, we see that to prove the upper bound in (A.1), it suffices to show that

$$\sum_{\sigma\in S_n:C_2=m} m!2^m\prod_{k\neq 2}2^{C_k}\ll n\times n! \tag{A.4}$$

for each $0\leq m\leq n/2$.

To establish (A.4), we use a double counting argument[10] as follows. For each permutation σ with exactly m 2-cycles, we assign a quantity $F(\sigma)$ which is the product of the number of ways to write down the 2-cycles of σ (counting ordering) and the number of ways to select some union E of the k-cycles of σ with $k\neq 2$.

If the 2-cycles of σ are $(x_1y_1),\dots,(x_my_m)$, then there are $m!2^m$ ways to write them down (counting all permutations in S_m and the permutations between x_j and y_j). Furthermore, there are $\prod_{k\neq 2}2^{C_k}$ ways to select E, which is a σ-invariant set disjoint from the $x_1,\dots,x_m,y_1,\dots,y_m$. This set E has some cardinality j between 0 and $n-2m$. Therefore,

$$\sum_{\sigma\in S_n:C_2=m} m!2^m\prod_{k\neq 2}2^{C_k}=\sum_{\sigma\in S_n:C_2=m}F(\sigma).$$

[10] One could in fact obtain much more precise asymptotics on (A.4) using the method of generating functions, but we will not need to do so here.

On the other hand, there are $\frac{n!}{(n-2m)!}$ ways to select $2m$ ordered elements $x_1, \ldots, x_m, y_1, \ldots, y_m$ of $\{1, \ldots, n\}$. For each j, there are then $\binom{n-2m}{j}$ ways to select E, and then to specify σ on E and on the complement of $E \cup \{x_1, \ldots, x_m, y_1, \ldots, y_m\}$ there are at most $j!(n-2m-j)!$ possibilities. (Notice that σ restricted to E is a permutation on E.) Putting all this together, we may bound the left-hand side of (A.4) by

$$\sum_{j=0}^{n-2m} \frac{n!}{(n-2m)!} \binom{n-2m}{j} j!(n-2m-j)! = \sum_{j=0}^{n-2m} n!$$

and the claim follows.

Now we turn to the lower bound. From Stirling's formula we have

$$\frac{2C_2!}{C_2!2^{C_2}} \gg C_2!2^{C_2}/\sqrt{C_2+1}$$

so by (A.3) (and crudely bounding $\sqrt{C_2+1}$ by $O(\sqrt{n})$) it suffices to show that

$$\sum_{\sigma \in S_n} C_2!2^{C_2} \prod_{k \neq 2} 2^{C_k} \gg n^2 \times n!$$

For this, it suffices to prove the matching lower bound

$$\sum_{\sigma \in S_n : C_2 = m} m!2^m \prod_{k \neq 2} 2^{C_k} \gg n \times n! \tag{A.5}$$

to (A.4) for each $0 \leq m \leq n/4$ (say).

We use the same double-counting argument as before. We write the left-hand side of (A.5) as $\sum_{\sigma \in S_n : C_2 = m} F(\sigma)$. We use the classical fact that as $n \to \infty$, the random variables $C_1, \ldots, C_k$ for any fixed k converge jointly to independent Poisson variables of intensities $1/1, \ldots, 1/k$ respectively (see e.g. [1]), so a positive constant fraction of S_n is 2-cycle free for $n \geq 0$. After fixing $x_1, \ldots, x_m, y_1, \ldots, y_m, E$, notice that any 2-cycle free permutation on E and on its complement will give a contribution to (A.5). Thus, we obtain a lower bound of the form

$$\gg \sum_{j=0}^{n-2m} \frac{n!}{(n-2m)!} \binom{n-2m}{j} j!(n-2m-j)! = n!(n-2m+1),$$

concluding the proof.

Now we consider the second moment for the GUE case. There are three differences here. Firstly, the factor of 2^{C_1} that was present in the GOE analysis (which arose from the fact that the diagonal entries had variance 2 instead of 1) is now absent. Secondly (and most importantly), in order for $\mathbf{E} I_\sigma \overline{I_\rho}$ to be non-vanishing, each cycle γ of length at least three in σ must appear in ρ also; the appearance of the inverse cycle γ^{-1} now leads to cancellation, in contrast with the GOE case. As such, the 2^{C_k} factors for $k \geq 3$ are also absent. Finally, the factor of 3^c in the above analysis becomes 2^c, due to the smaller value of the fourth moment $\mathbf{E}|\zeta_{ij}|^4$ of the off-diagonal entries in the GUE case. Repeating the GOE arguments, one reduces to showing that

$$\sum_{\sigma \in S_n : C_2 = m} m!2^m \ll n!$$

for all $0 \leq m \leq n/2$, and

$$\sum_{\sigma \in S_n : C_2 = m} m! 2^m \gg n!$$

for all $0 \leq m \leq n/4$. But this can be achieved by a routine modification of the above arguments (with the role of the additional set E, which represented the $\prod_{k \neq 2} 2^{C_k}$ factor, now omitted).

Remark 17. An inspection of the above argument shows that the hypotheses that M_n are distributed according to GOE or GUE can be relaxed to the assertion that M_n matches GOE or GUE to fourth order off the diagonal and to second order on the diagonal.

References

[1] R. Arratia, S. Tavaré, The cycle structure of random permutations, Ann. Probab. 20 (3) (1992) 1567–1591.
[2] Z. Bai, J. Silverstein, Spectral Analysis of Large Dimensional Random Matrices, Science Press, 2006.
[3] J. Bourgain, V. Vu, P.M. Wood, On the singularity probability of discrete random matrices, J. Funct. Anal. 258 (2) (2010) 559–603.
[4] E. Brézin, S. Hikami, Characteristic polynomials of random matrices, Comm. Math. Phys. 214 (1) (2000) 111–135.
[5] B.M. Brown, Martingale central limit theorems, Ann. Math. Statist. 42 (1) (1971) 59–66.
[6] K. Costello, T. Tao, V. Vu, Random symmetric matrices are alsmot surely singular, Duke Math. J. 135 (2006) 395–413.
[7] R. Delannay, G. Le Caer, Distribution of the determinant of a random real-symmetric matrix from the Gaussian orthogonal ensemble, Phys. Rev. E 62 (2000) 1526–1536.
[8] A. Dembo, On random determinants, Quart. Appl. Math. 47 (2) (1989) 185–195.
[9] L. Erdős, B. Schlein, H.-T. Yau, Wegner estimate and level repulsion for Wigner random matrices, Int. Math. Res. Not. 2010 (2010) 436–479.
[10] L. Erdős, H.-T. Yau, J. Yin, Bulk universality for generalized Wigner matrices, arXiv:1001.3453.
[11] P.J. Forrester, N.E. Frankel, Applications and generalizations of Fisher–Hartwig asymptotics, J. Math. Phys. 45 (2004) 2003–2028.
[12] G.E. Forsythe, J.W. Tukey, The extent of n random unit vectors, Bull. Amer. Math. Soc. 58 (1952) 502.
[13] T. Garoni, On the asymptotics of some large Hankel determinants generated by Fisher–Hartwig symbols defined on the real line, J. Math. Phys. 46 (4) (2005) 043516.
[14] V.L. Girko, The central limit theorem for random determinants, Theory Probab. Appl. 24 (4) (1979) 729–740 (in Russian). Translation.
[15] V.L. Girko, Theory of Random Determinants, Kluwer Academic Publishers, 1990.
[16] V.L. Girko, A Refinement of the central limit theorem for random determinants, Theory Probab. Appl. 42 (1) (1998) 121–129 (in Russian); translation.
[17] N.R. Goodman, Distribution of the determinant of a complex Wishart distributed matrix, Ann. Statist. 34 (1963) 178–180.
[18] J. Kahn, J. Komlós, E. Szemerédi, On the probability that a random ± 1 matrix is singular, J. Amer. Math. Soc. 8 (1995) 223–240.
[19] R. Killip, Gaussian fluctuations for β ensembles, Int. Math. Res. Not. (8) (2008) 19. Art. ID rnn007.
[20] A. Knowles, J. Yin, Eigenvector distribution of Wigner matrices, arXiv:1102.0057.
[21] J. Komlós, On the determinant of (0, 1) matrices, Studia Sci. Math. Hungar. 2 (1967) 7–21.
[22] J. Komlós, On the determinant of random matrices, Studia Sci. Math. Hungar. 3 (1968) 387–399.
[23] I.V. Krasovsky, Correlations of the characteristic polynomials in the Gaussian unitary ensemble or a singular Hankel determinant, Duke Math. J. 139 (3) (2007) 581–619.
[24] A. Maltsev, B. Schlein, A Wegner estimate for Wigner matrices, arXiv:1103.1473.
[25] H. Nguyen, On the least singular value of random symmetric matrices, arXiv:1102.1476.
[26] H. Nguyen, V. Vu, Random matrix: law of the determinant, arXiv:1112.0752 (submitted for publication).
[27] H. Nyquist, S.O. Rice, J. Riordan, The distribution of random determinants, Quart. Appl. Math. 12 (1954) 97–104.
[28] A. Prékopa, On random determinants I, Studia Sci. Math. Hungar. 2 (1967) 125–132.
[29] G. Rempala, J. Wesolowski, Asymptotics for products of independent sums with an application to Wishart determinants, Statist. Probab. Lett. 74 (2) (2005) 129–138.

[30] A. Rouault, Asymptotic behavior of random determinants in the Laguerre, Gram and Jacobi ensembles, ALEA Lat. Am. J. Probab. Math. Stat. 3 (2007) 181–230.
[31] G. Szekeres, P. Turán, On an extremal problem in the theory of determinants, Math. Term. Tud. Értesitö 55 (1937) 796–805.
[32] T. Tao, V. Vu, On random ± 1 matrices: singularity and determinant, Random Structures Algorithms 28 (2006) 1–23.
[33] T. Tao, V. Vu, On the singularity probability of random Bernoulli matrices, J. Amer. Math. Soc. 20 (2007) 603–628.
[34] T. Tao, V. Vu, Random matrices: universality of local eigenvalue statistics up to the edge, Comm. Math. Phys. 298 (2) (2010) 549–572.
[35] T. Tao, V. Vu, Random matrices: universality of the local eigenvalue statistics, Acta Math. 206 (2011) 127–204.
[36] T. Tao, V. Vu, Random covariance matrices: university of local statistics of eigenvalues, Ann. Probab. 40 (2012) 1283–1315.
[37] T. Tao, V. Vu, Random matrices: universal properties of eigenvectors, Random matrices: Theory and Applications, arXiv:1103.2801 (in press).
[38] H. Trotter, Eigenvalue distributions of large Hermitian matrices: Wigner's semicircle law and a theorem of Kac, Murdock, and Szegö, Adv. Math. 54 (1) (1984) 67–82.
[39] P. Turán, On a problem in the theory of determinants, Acta Math. Sinica 5 (1955) 41l–423 (in Chinese).
[40] R. Vershynin, Invertibility of symmetric random matrices, arXiv:1102.0303.